PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The office work and stretch training (OST) study: Effects on the prevalence of musculoskeletal diseases and gender differences - A non-randomized control study.
AUTHORS	Holzgreve, Fabian; Maltry, Laura; Hänel, Jasmin; Schmidt, Helmut; Bader, Andreas; Frei, Markus; Groneberg, David; Ohlendorf, Daniela; van Mark, Anke

VERSION 1 – REVIEW

REVIEWER	Kuptniratsaikul, V. Mahidol University, Bangkok
REVIEW RETURNED	16-Nov-2020

GENERAL COMMENTS	November 16, 2020 Dear Editor of BMJ Open Thank you for letting me review the article entitled “The office work and stretch training (OST) study: Effects on the prevalence of musculoskeletal diseases and gender differences”. After completely review this article, I have some comments as followings: Abstract • How long did it take for one training unit on the "five-Business "device? Authors better clarify in method part of abstract too. Introduction • Please clarify "five-Business" training to enhance more understanding what authors want to do. Method • Did authors calculate for sample size? • Did authors include gender differences into the formula for SSC? • What method is used to allocate participants? • How long for treatment duration for each session? Result • Authors stated that the duration of training was 3 months. And authors measure NQ twice; before and at the end of 3 months. If so, how can authors get the 12-month and lifelong prevalence of MSD complaints? • Did authors record adverse events from attending the "five-Business" device? • What did MSE stand for (page 10, line 244) or it's just typo error? Discussion • Authors did not report the days and main causes of sick leave in the result section. • Did participants have to have trainer for each training session? In addition, this program need the "five-Business" device. So, it could be a limitation concerning generalizability.

REVIEWER	Hoe, Victor C. W. University of Malaya, Department of Social and Preventive Medicine
REVIEW RETURNED	25-Nov-2020

GENERAL COMMENTS	Thank you for the opportunity to review the report titled "The office work and stretch training (OST) study: Effects on the prevalence of musculoskeletal diseases and gender differences". The manuscript reports on OST intervention programme, which consists of the individualised stretching exercise of duration 20 seconds each time for at least twice in a week. The study period was for 12 weeks, with data collected one week before and after the completion of the intervention. The main outcome of interest was reporting of 12-month prevalence of musculoskeletal pain using the Nordic Questionnaire for musculoskeletal complaints. The study found that there was a reduction of complaints in both the intervention and control groups. The was also a significant difference comparing the intervention and control groups. The report consists of some major issues that need to be addressed. 1. The selection of the participants was based non-probability sampling technique. The participants were all volunteers, this is prone for selection bias. The results for the baseline prevalence does not fulfil external validity and should not be included as part of the discussion. 2. The allocation of the intervention and control group was based on the participant preference, this has led to an unequal size for the intervention and control group. The use of non-randomised allocation of participants for the intervention and control group will lead to allocation bias. In a non-RCT study, the two groups should be comparable. In the study, the information on the similarity or differences between the two groups was presented. The participants who have opted for the intervention group are more inclined to ensure the success of the study. Although this has been listed as one of the limitations and the reason for choosing a non-RCT approach has also been discussed, the improvement from the intervention from a non-RCT study can be due to factors other than the intervention. Further, in the drop-out from the intervention group is also very large, i.e., 27%. 3. The main outcome variable assessed was the 12-month prevalence of musculoskeletal pain. Since the study period was only for 12 weeks, and the data collection period was on the week before and after the study, using a 12-month prevalence will not be appropriate, as the period covered the time before the intervention has started. This means the two measures have a intersect of 40 weeks. The information of the 12-month prevalence post-intervention is not a reliable measure when the intervention and data collection was only for a period of 12-weeks.

	Suggestion on how to improve the report. 1. Since this is a non-RCT study and the number of participants in the intervention group is not similar to the control group, it would not be appropriate to compare the two groups. One option is to report the study as a before and after study and focus on the issue that contributes to the success of the reduction in reporting of musculoskeletal pain. 2. Since t

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. V. Kuptniratsaikul, Mahidol University, Bangkok Comments to the Author:
November 16, 2020
Dear Editor of BMJ Open
Thank you for letting me review the article entitled “The office work and stretch training (OST) study: Effects on the prevalence of musculoskeletal diseases and gender differences”. After completely review this article, I have some comments as followings:

Abstract
• How long did it take for one training unit on the "five-Business "device? Authors better clarify in method part of abstract too.
One Training unit had a duration of about 10 minutes. We added this information in the abstract (line 46).
Introduction
• Please clarify "five-Business" training to enhance more understanding what authors want to do.
Thank you for the advice. We added information about the five-business program to the introduction (line 118-122).
Method
• Did authors calculate for sample size?
The sample size was calculated prior to the study. The calculations have been published in the methodology article (Holzgreve F, Maltry L, Lampe J, et al. The office work and stretch training (OST) study: an individualized and standardized approach for reducing musculoskeletal disorders in office workers. Journal of Occupational Medicine and Toxicology 2018;13(1):37. doi: 10.1186/s12995-018-0220-y). We added the term “The sample size was calculated prior to the study. The calculations have been published in the related methodology article 27.” in line 211-213 in the “Statistical analysis” section.
• Did authors include gender differences into the formula for SSC?
We added the term: “Gender differences have not been included in the calculation of the sample size.” line 211-213 in the “Statistical analysis” section.
• What method is used to allocate participants?
We added to the method section (134-144): “Prior to the study we randomized in which department of the factory the intervention will take place. The following recruitment strategy consisted of an internal e-mail which was sent by the health department. Via an integrated link, employees could register for participation on a voluntary basis (Fig. 1); the registration deadline was set at two weeks. It was communicated that the training would be carried out during working hours and all participants were asked to pursue (only) their usual leisure activities during the intervention period. In close cooperation with the works council, the training was carried out during working hours. In order to provide all employees with the same opportunity to participate in an intervention during work hours, a waiting

control group was included as part of the study design. In addition, a non-randomized allocation procedure based on availability was conducted to enable every employee to participate."
Furthermore, we extended the limitation segment (line 325-336): "When interpreting the presented results, the lack of randomization has to be taken into account. A randomized allocation to the study groups was not possible due to the organizational structure of the company. However, the aim of this study was to investigate the effects of a WHP program deliberately in the field. In our opinion, this also represents the strength of this study. It was carried out under realistic working conditions with the involvement of the works council within the framework of the employees' working hours on a voluntary basis. Furthermore, such WHP programs are usually applied in large companies, where cooperation with the works council is essential. Any other approach would have been simply not being possible in this case. In addition, subjects had to arrange training in accordance with business appointments and holidays, randomization would have threatened the feasibility of this study. This might also explain differences in the group size and gender distribution between the intervention and control group."

• How long for treatment duration for each session?
The treatment duration was about 10 minutes for each session. We added "One training unit had a duration of about 10 minutes." in line 186.
Result
• Authors stated that the duration of training was 3 months. And authors measure NQ twice; before and at the end of 3 months. If so, how can authors get the 12-month and lifelong prevalence of MSD complaints?
In order to evaluate musculoskeletal complaints, we chose to use the Nordic Questionnaire, which had been applied internationally in various studies. In order to ensure the highest possible level of standardization, we did not adapt the Nordic Questionnaire to the treatment period.
• Did authors record adverse events from attending the "five-Business" device?
Adverse events from attending the treatment were almost none. We added the term "Adverse events due to treatment were almost none. One subject terminated participation due to thoracic spine discomfort." (line 225-227) to the results section.
• What did MSE stand for (page 10, line 244) or it's just typo error?
Thank you for the hint. This was just a typo error, which was corrected. It should have been MSD.
Discussion
• Authors did not report the days and main causes of sick leave in the result section.
We added an explanation to the discussion section (line 339-340). In response to the intervention of the works council, we were not allowed to evaluated the days and main causes of sick leave.
• Did participants have to have trainer for each training session? In addition, this program need the "five-Business" device. So, it could be a limitation concerning generalizability.
Training on this device can be done by yourself after a one-time instruction. For the application of the stretching concept "FIVE" special devices were designed. Typically, there is one device for each stretching exercise. For the implementation of this concept as a workplace health promotion program, the manufacturer designed one device, were all the main stretching exercises for the trunk can be conducted. In general, this concept always relies on its suitable stretching devices.
We added (line 342-344) an explanation to the discussion/limitations.
Reviewer: 2
Dear Dr. Victor C. W. Hoe,
Thank you for the detailed review. We have taken your criticism and tried to implement your points without changing the whole structure of the manuscript. We are aware that methodological errors such as the use of a non-probability sampling technique or the lack of randomization have led to a reduced internal validity. However, the aim of this study was to investigate the effects of a WHP program deliberately in the field. In our opinion, this also represents the strength of this study. It was carried out under realistic working conditions with the involvement of the works council within the framework of the employees' working hours on a voluntary basis. This methodological approach aims to test the effects of such an intervention under application conditions that are as realistic as possible.

Accordingly, the study has a high external validity. For a prior allocation in the sense of a randomized sampling technique or a prior randomization, consent from the works council is difficult to obtain. However, such WHP programs are usually applied in large companies, where cooperation with the works council is essential. Any other approach would have been simply not being possible in this case. We think that methodological difficulties of this kind are not uncommon in workplace health promotion and that these circumstances must be taken into account.
Below we comment on each of your points.
Dr. Victor C. W. Hoe, University of Malaya Comments to the Author:
Thank you for the opportunity to review the report titled "The office work and stretch training (OST) study: Effects on the prevalence of musculoskeletal diseases and gender differences".

The manuscript reports on OST intervention programme, which consists of the individualised stretching exercise of duration 20 seconds each time for at least twice in a week. The study period was for 12 weeks, with data collected one week before and after the completion of the intervention. The main outcome of interest was reporting of 12-month prevalence of musculoskeletal pain using the Nordic Questionnaire for musculoskeletal complaints. The study found that there was a reduction of complaints in both the intervention and control groups. The was also a significant difference comparing the intervention and control groups.

The report consists of some major issues that need to be addressed.

1. The selection of the participants was based non-probability sampling technique. The participants were all volunteers, this is prone for selection bias. The results for the baseline prevalence does not fulfil external validity and should not be included as part of the discussion.
Thank you for your comment. We were not allowed to randomize selection among participants because of the intervention by the works council and legal department and their concerns about data protection. We implemented this term in line 336-337.
2. The allocation of the intervention and control group was based on the participant preference, this has led to an unequal size for the intervention and control group. The use of non-randomised allocation of participants for the intervention and control group will lead to allocation bias.
In the methodology article, which was published 2018 (Holzgreve F, Maltry L, Lampe J, et al. The office work and stretch training (OST) study: an individualized and standardized approach for reducing musculoskeletal disorders in office workers. Journal of Occupational Medicine and Toxicology 2018;13(1):37. doi: 10.1186/s12995-018-0220-y), we published a power analysis for this study. The calculated group sizes were not equal - 250 in the intervention group and 100 in the control group. Unfortunately, despite a non-randomized distribution, we had high dropouts, resulting in a somewhat smaller group of successful intervention participants.
We extended the methods section (line 134-144) and the discussion (325-335) in order to explain and clarify the chosen study design.
We share your concerns and we are aware that this study has indeed a selection bias, since the allocation of the subjects was non-randomized. The aim of this study was to investigate the effect of this WHP program on MSD in office workers. The works council had approved the implementation of this study only on the condition that the training was conducted within working hours and payment for the training time, with respect of data protection and privacy of participants so randomization was not allowed. In order to provide all employees with the same opportunity to participate in an intervention during work hours, a waiting control group was included as part of the study design.
In a non-RCT study, the two groups should be comparable. In the study, the information on the similarity or differences between the two groups was presented.
Please see our argumentations and new implementations on point 2.
The participants who have opted for the intervention group are more inclined to ensure the success of the study. Although this has been listed as one of the limitations and the reason for choosing a non-

RCT approach has also been discussed, the improvement from the intervention from a non-RCT study can be due to factors other than the intervention.
Thank you for comment. In line 136-138 we have stated that participants were asked to pursue their usual leisure activities, in order to control for confounders. Furthermore, we added the term (line 194-198): “Prior to the study, participants were instructed not to start any new treatments during the intervention period. This did not apply to necessary treatments. In order to control for confounders each participant had to fill out a sports diary on every appointment. If new treatments were started within the intervention period that had an impact on the musculoskeletal system, this resulted in study exclusion.”
Further, in the drop-out from the intervention group is also very large, i.e., 27%.
This is true. However, it should be noted at this point that the participation requirements were challenging in the context of full employment. Regular study participation over three months could not be met due to the reasons mentioned in Figure 1. Prior to the study, it was clearly communicated in consultation with the management that professional reasons take precedence at all times.
3. The main outcome variable assessed was the 12-month prevalence of musculoskeletal pain. Since the study period was only for 12 weeks, and the data collection period was on the week before and after the study, using a 12-month prevalence will not be appropriate, as the period covered the time before the intervention has started. This means the two measures have a intersect of 40 weeks. The information of the 12-month prevalence post-intervention is not a reliable measure when the intervention and data collection was only for a period of 12-weeks.
In order to evaluate musculoskeletal complaints, we chose to use the Nordic Questionnaire, which had been applied internationally in various studies. In order to ensure the highest possible level of standardization, we did not adapt the Nordic Questionnaire to the treatment period.

Suggestion on how to improve the report.

1. Since this is a non-RCT study and the number of participants in the intervention group is not similar to the control group, it would not be appropriate to compare the two groups. One option is to report the study as a before and after study and focus on the issue that contributes to the success of the reduction in reporting of musculoskeletal pain.
According to the power analysis conducted by the Institute of Biostatistics of our University published 2018 (Holzgreve F, Maltry L, Lampe J, et al. The office work and stretch training (OST) study: an individualized and standardized approach for reducing musculoskeletal disorders in office workers. Journal of Occupational Medicine and Toxicology 2018;13(1):37. doi: 10.1186/s12995-018-0220-y) the calculated group sized were not equal for this non-RCT study. Therefore, we never aimed to reach equal group sizes.

2. Since t

VERSION 2 – REVIEW

REVIEWER	Hoe, Victor C. W. University of Malaya, Department of Social and Preventive Medicine
REVIEW RETURNED	01-Mar-2021

GENERAL COMMENTS	Thank you for the response to my earlier comments. The explanation and the changes made to the article is acceptable and appropriate.